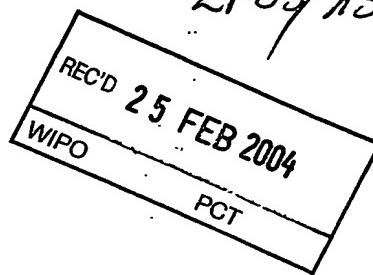


EPO 3 / 13964



**PRIORITY  
DOCUMENT**

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

**Prioritätsbescheinigung über die Einreichung  
einer Patentanmeldung**

**Aktenzeichen:** 102 57 479.0

**Anmeldetag:** 9. Dezember 2002

**Anmelder/Inhaber:** Roche Diagnostics GmbH, Mannheim/DE

**Bezeichnung:** Optimierte Proteinsynthese

**IPC:** C 12 N 15/63

**Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.**

München, den 23. Oktober 2003  
Deutsches Patent- und Markenamt  
Der Präsident  
Im Auftrag

A handwritten signature in black ink, appearing to read "Weber". Below the signature, the name "Weber" is printed in a smaller font.

# WEICKMANN & WEICKMANN

Patentanwälte  
European Patent Attorneys · European Trademark Attorneys

DIPL.-ING. H. WEICKMANN (bis 31.1.01)  
DIPL.-ING. F. A. WEICKMANN  
DIPL.-CHEM. B. HUBER  
DR.-ING. H. LISKA  
DIPL.-PHYS. DR. J. PRECHTEL  
DIPL.-CHEM. DR. B. BÖHM  
DIPL.-CHEM. DR. W. WEISS  
DIPL.-PHYS. DR. J. TIESMEYER  
DIPL.-PHYS. DR. M. HERZOG  
DIPL.-PHYS. B. RUTTENSPERGER  
DIPL.-PHYS. DR. ING. V. JORDAN  
DIPL.-CHEM. DR. M. DEY  
DIPL.-FORSTW. DR. J. LACHNIT

Unser Zeichen:  
**29415P DE/WWpu**

Anmelder:  
**Roche Diagnostics GmbH**  
**Sandhofer Str. 112 - 132**  
**68305 Mannheim**

---

**Optimierte Proteinsynthese**

---

## Optimierte Proteinsynthese

### Beschreibung

5

Die Erfindung betrifft ein Verfahren zur optimierten Herstellung von Proteinen in einem *in vitro* oder *in vivo* Expressionssystem sowie dafür geeignete Reagenzien.

- 10 Hannig, G. & Makrides, S.C. (1998) Tibtech Vol 16, pp 54-60, haben Strategien zur Optimierung der heterologen Proteinexpression in *E. coli* beschrieben. Ein entscheidender Faktor ist hierbei die Effizienz der Initiation der Translation. Die Verwendung bestimmter Codons spielt hierbei eine gewisse Rolle. So konnten George et al. (1985) DNA Vol 4, pp 273-281, zeigen, dass die Expression eines heterologen Genes gesteigert werden kann, wenn man in dem Bereich nach dem Startcodon solche Codons verwendet, die in *E. coli* Genen häufig benutzt werden. Besonders wichtig für die Translations-Initiation sind hauptsächlich strukturelle Elemente am 5'-Ende der mRNA. Von Makrides (1996) Microbiol. Rev. Vol 60, pp 512-20 538, wurden verschiedene Translations-Verstärker-Sequenzen, wie beispielsweise eine Sequenz aus dem T7-Phagen Gen10 Leader und eine U-reiche Sequenz aus der 5'-untranslatierten Region einiger mRNA's, wie beispielsweise des *E. coli* atpE Gens, beschrieben.
- 25 Bislang wurden keine universell einsetzbaren Translations-Initiationssequenzen beschrieben. Es wurden jedoch Strategien beschrieben, die das Potential der Sekundärstruktur-Bildung am 5'-Ende der mRNA reduzieren. Insbesondere wurde die Ribosomen-Bindungsstelle mit Adenin und Thymin Bausteinen angereichert. Stenstöm et al. (2001) Gene Vol 263, pp 273-284, zeigten, dass stark exprimierte *E. coli* Gene insbesondere bei dem dem Startcodon folgenden + 2 Codon einen hohen
- 30

Gehalt an Adeninen aufweisen. Allerdings gibt es auch für diese Regel viele positive und negative Ausnahmen.

Schließlich zeigten Pederson-Lane et al. (1997) Protein Expr. Purif. Vol 10, 5 pp 256-262, dass ein hoher GC-Gehalt unmittelbar nach dem Startcodon negativ für die Expression ist und mit Konversion der Purin Basen des 3., 4. und 5. Codons gegen Thymidin Basen die Expression der Thymidylatsynthase auf 25 % des Gesamtproteins gesteigert werden konnte.

10

Es wird angenommen, dass bei all diesen Maßnahmen die Zugängigkeit der 30S Ribosomenuntereinheit zu der messenger RNA eine entscheidende Rolle spielt. Besonders wichtig ist hierbei der freie Kontakt an die nach Shine und Dalgarno benannte Sequenz unmittelbar vor dem Startcodon und 15 der Kontakt zu dem Startcodon selbst. Sind hingegen diese Sequenzelemente in stabilen RNA-Sekundärstrukturen gebunden, so läuft die Initiation der Translation sehr ineffizient. Tessier et al. (1984) Nucl. Ac. Res. Vol 12, pp 7663-7675, zeigten in einer systematischen Untersuchung, dass durch eine gezielte Mutation diese Form von 20 Stämmchen und Schleifen (sogenannten Stem-Loops oder Hairpin-Loops) ähnelnden Sekundärstrukturen aufgelöst und damit die Effizienz der Translation erheblich gesteigert werden kann. Der Effekt dieser Sekundärstrukturen auf die Translation kann nach deren thermodynamischen Parametern berechnet werden. So bewirkt eine 25 Stabilisierung um 1,4 kcal/mol eine 10fache Reduzierung der Expression (Gold (1988) Ann. Rev. Biochem, Vol 57, pp 199-233) und eine Stabilisierung um 2,3 kcal/mol reduziert die Bindung des Ribosoms um eine Größenordnung (de Smit & van Duin (1994) J. Mol. Biol. Vol 244, pp 144-150).

30

Von Sprengart et al. (1996) EMBO Vol 15, pp 665-674, wurde als weiterer Translations-Verstärker die sogenannte „Downstream Box“, ein

Sequenzelement direkt nach dem Startcodon der T7-Gene mit Homologie zur ribosomalen 16S RNA beschrieben. Es wird angenommen, dass dieses Element durch eine Wechselwirkung der beiden homologen Basenpaare die Bindung der 30S Ribosomenuntereinheit verstärkt. Allerdings eignet sich auch dieses Element nicht als universeller Translations-Verstärker.

Die Nachteile bekannter Verfahren liegen darin, dass für jedes neue Gen eine Optimierung der 5'-Region der mRNA entweder in der 5'-untranslatierten Region oder in der translatierten Region durchgeführt werden muss, um die Codon-Verwendung zu optimieren, oder unerwünschte Sekundärstrukturen der mRNA mit Auswirkung auf die Shine-Dalgarno Sequenz oder das Startcodon zu vermeiden. Dies bedarf im Regelfall einer aufwändigen Analyse der RNA-Struktur mit entsprechenden Programmen (z.B. Mukund et al. (1999) Curr. Science Vol 76, pp 1486-1490, oder Jaeger et al. (1990) Meth. Enzymol. Vol 183, pp 281-306), sowie mehrerer PCR-Amplifikationen und Klonierungsschritte. Will man auf diese Weise eine große Anzahl von Genen, beispielweise aus einer Genbank exprimieren, so ist dazu in jedem Falle eine genaue Kenntnis der Sequenz erforderlich, weshalb diese Methoden nicht auf unbekannte Gene angewendet werden können. Selbst bei Kenntnis der Sequenzen wäre dieses Verfahren wesentlich aufwändiger als eine universell anwendbare Methode.

Ein weiterer Ansatz zur Verstärkung der Translation besteht darin, als universellen Translations-Enhancer ein Fusionsprotein mit einem stark exprimierten Gen zu bilden, an dessen C-terminales Ende das gewünschte Gen gesetzt wird. Als Beispiel für den Erfolg dieser Strategie ist die Fusion mit dem Ubiquitin Gen zu nennen, wie sie von Butt et al. (1989) PNAS Vol 86, pp 2540-2544, durchgeführt wurde.

30

Auch dieser Ansatz kann jedoch nicht ohne Weiteres auf die Expression beliebiger Gene übertragen werden. Werden nämlich Fusionsproteine

verwendet, so wird am N-Terminus des Proteins eine mehr oder weniger große Fusion angefügt, die aufgrund der Größe und Eigenschaften des Fusionspartners mit der Funktion des gewünschten Proteins interferieren können. Je kleiner die Fusionsproteine oder Teile von diesen gewählt werden, desto geringer ist in vielen Fällen auch deren translationssteigernde Wirkung. Bei großen Fusionsproteinen zeigt sich ein weiterer Nachteil prokaryontischer Expressionssysteme: Es steigt gleichzeitig die Wahrscheinlichkeit für unvollständige Transkription oder Translation durch vorzeitige Termination oder interne Initialisierung. Auch die Wahrscheinlichkeit für proteolytischen Abbau ist erhöht.

Es besteht somit ein Bedürfnis, ein Verfahren zur optimierten Herstellung von Proteinen bereitzustellen, bei dem die Nachteile des Standes der Technik zumindest teilweise beseitigt sind.

15

Ein Gegenstand der Erfindung ist ein Verfahren zur Herstellung eines Proteins, umfassend die Schritte:

20

(a) Bereitstellen einer für das Protein kodierenden Nukleinsäuresequenz, wobei 3'-seitig des Translations-Startcodons eine heterologe Nukleinsäuresequenz im korrekten Leseraster eingefügt wird, die so gewählt wird, dass in einem Abstand von 6-30 Nukleotiden 3'-seitig des Translations-Startcodons eine Stem-Loop-Struktur ausgebildet wird,

25

(b) Bereitstellen eines zur Expression des Proteins geeigneten Expressionssystems und

(c) Einbringen der Nukleinsäuresequenz gemäß (a) in das Expressionssystem gemäß (b) unter Bedingungen, dass das Protein synthetisiert wird.

30

Die erfindungsgemäße Lösung für ein universell optimiertes Expressionskonstrukt besteht darin, dass ein kleines heterologes DNA-Sequenzelement mit vorzugsweise maximal 201 Basenpaaren,

- besonders bevorzugt maximal 45 Basenpaaren, unmittelbar nach dem Startcodon des zu exprimierenden Gens eingefügt wird, welches die Ausbildung stabiler Stem-Loop-Strukturen im Bereich der Shine-Dalgarno-Sequenz und des Startcodons weitgehend verhindert und dadurch zur optimierten Translations-Initiation und optimierter Proteinsynthese führt. Es wird somit ein Fusionsprotein gebildet, wobei vorzugsweise nur ein kleines Peptid mit maximal 67 Aminosäuren und besonders bevorzugt maximal 15 Aminosäuren an das gewünschte Protein angefügt wird.
- 10 Eine wichtige Voraussetzung für das heterologe DNA-Sequenzelement ist, dass es im korrekten Leseraster eingefügt wird, d.h. dass keine Rasterverschiebung im zu exprimierenden Gen stattfindet. Eine weitere wichtige Eigenschaft des heterologen DNA-Sequenzelements ist, dass sich in der transkribierten RNA in einem Abstand von 6-30 Basen, vorzugsweise 15 12-21 Basen, hinter dem Startcodon eine stabile Stem-Loop-Struktur ausbilden kann, wobei die Basenpaarung in der Stem-Loop-Struktur zumindest teilweise durch die eingefügte Sequenz bewirkt wird. Diese Stem-Loop-Struktur soll so beschaffen sein, dass sie nach der erfolgten Initiation der Translation durch das Ribosom wieder aufgelöst werden kann und somit nicht zu einem Abbruch der Translation führt. Diese durch Einführung der heterologen Nukleinsäuresequenz in das Expressionskonstrukt entstandene Stem-Loop-Struktur kann sich bei nahezu jedem Gen in der gleichen Weise ausbilden und dadurch verhindern, dass die für die Translations-Initiation wichtigen Sequenzen vor dem Loop 20 25 größere Sekundärstrukturen mit der kodierenden Sequenz des Gens eingehen können. Der Bereich unmittelbar vor dieser Stem-Loop-Struktur und nach dem Startcodon ist vorzugsweise eine Sequenz ohne Sekundärstruktur und kann auch keine Sekundärstruktur mit der 5'-untranslatierten Region ausbilden. Besonders bevorzugt ist in dieser 30 Region eine GC-arme Sequenz, da bei einer derartigen Sequenz die Ausbildung stabiler Sekundärstrukturen mit Sequenzen innerhalb der translatierten Region minimiert wird.

Das heterologe Nukleinsäure-Sequenzelement kann in die Zielsequenz, z.B. in einen Plasmidvektor zur Expression von heterologen Genen mit bekannten Klonierungs- oder/und Amplifikationstechniken eingefügt werden. Möglich ist z.B. der Aufbau dieser Sequenz durch PCR-Primer zur 5 Klonierung des gewünschten Gens oder durch Primer, mit denen DNA-Expressionskonstrukte für die *in vitro* Proteinexpression hergestellt werden.

Das erfindungsgemäße Verfahren kann zur Herstellung und gegebenenfalls 10 Gewinnung von Proteinen in *in vitro* Expressionssystemen eingesetzt werden. Beispiele für geeignete *in vitro* Expressionssysteme sind prokaryontische *in vitro* Expressionssysteme, wie etwa Lysate von gram-negativen Bakterien, beispielsweise von *Escherichia coli*, oder gram-positiven Bakterien, wie beispielsweise *Bacillus subtilis*, oder eukaryontische *in vitro* Expressionssysteme, wie etwa Lysate von 15 Säugerzellen, wie beispielsweise von Kaninchen, Reticulocyten, humanen Tumorzelllinien, Hamsterzelllinien, oder anderen Wirbeltierzellen, wie beispielsweise Oozyten und Eiern von Fischen und Amphibien, sowie Insektenzelllinien, Hefezellen, Algenzellen oder Extrakte aus Pflanzenkeimen.

20 Alternativ kann die Herstellung des Proteins in einem *in vivo* Expressionssystem erfolgen, wobei eine prokaryontische Zelle, z.B. eine gram-negative prokaryontische Wirtszelle, insbesondere eine *E. coli* Zelle, oder eine gram-positive prokaryontische Zelle, insbesondere eine *Bacillus subtilis* Zelle, eine eukaryontische Wirtszelle, z.B. eine Hefezelle, eine 25 Insektenzelle oder eine Wirbeltierzelle, insbesondere eine Amphibien-, Fisch-, Vogel- oder Säugerzelle, oder ein nicht-humaner eukaryontischer Wirtsorganismus als Expressionssystem verwendet werden kann.

30 Die Einführung der heterologen Nukleinsäure-Sequenz in die für das gewünschte Protein kodierenden Nukleinsäure-Sequenz kann durch Standardmethoden der Molekularbiologie, z.B. durch Klonierung, wie etwa

Restriktionsspaltung oder/und Ligation, durch Rekombination oder/und durch Nukleinsäureamplifikation erfolgen. Die Nukleinsäure-Zielsequenz kann dabei auf einem geeigneten Vektor, z.B. einem Plasmidvektor zur Expression von heterologen Genen, oder auf einem Konstrukt für eine *in vitro* Proteinexpression vorliegen. Besonders bevorzugt ist eine Nukleinsäureamplifikation in einem oder mehreren Schritten, wobei durch Auswahl geeigneter Primer die heterologe Nukleinsäuresequenz und gegebenenfalls Expressionskontrollsequenzen, wie etwa Promotor, ribosomale Bindungsstellen und Terminatoren, an die für das gewünschte Protein kodierende Nukleinsäure-Sequenz angefügt werden. Besonders bevorzugt ist eine Zweistufen-PCR, wobei in einer ersten Stufe zumindest ein Teil der heterologen Nukleinsäure-Sequenz an eine Nukleinsäure-Zielsequenz, die für das gewünschte Protein kodiert, und in einem zweiten Schritt Expressionskontrollsequenzen angefügt werden. Eine bevorzugte Ausführungsform zur Durchführung einer Zweischritt-PCR ist in den Beispielen exemplarisch dargestellt.

Die heterologe Nukleinsäure-Sequenz, die in der Lage ist, eine Stem-Loop-Struktur 3'-seitig des Translations-Startcodons auszubilden, wird im korrekten Leseraster 3'-seitig des Translations-Startcodons, üblicherweise des ersten ATG-Codons, in die für das gewünschte Protein kodierende Nukleinsäure-Sequenz eingefügt. Vóorzugsweise erfolgt die Einfügung in einem Abstand bis zu 6 Nukleotiden, besonders bevorzugt unmittelbar nach dem Translations-Startcodon. Eine Einfügung im "korrekten Leseraster" bedeutet dabei, dass keine Verschiebung des Leserahmen in der Protein-kodierenden Nukleinsäure-Sequenz erfolgt. Dies bedeutet wiederum, dass die Länge der heterologen Nukleinsäure-Sequenz in Nukleotiden ein Vielfaches von 3 ist. Bevorzugt ist die Länge im Bereich von 6-201 Nukleotiden, besonders bevorzugt im Bereich von 12-45 Nukleotiden.

30

Die Einfügung der heterologen Nukleinsäure-Sequenz in die Protein-kodierende Nukleinsäure-Sequenz erfolgt derart, dass eine Stem-Loop-

Struktur in einem geeigneten Abstand 3'-seitig des Translations-Codons ausgebildet wird. Der Abstand (zwischen dem letzten Nukleotid des Translationsstart-Codons und dem ersten Nukleotid des Stem) beträgt günstigerweise 6-30 Nukleotide, besonders bevorzugt 12-21 Nukleotide.

5 5'-seitig derjenigen Sequenzen, die für die Bildung der Stem-Loop-Struktur vorgesehen sind, enthält die heterologe Nukleinsäure-Sequenz vorzugsweise einen AT-reichen Bereich, d.h. einen Bereich mit einem AT-Gehalt >50 %, insbesondere >60 %.

10 Die Länge des Stems in der Stem-Loop-Struktur liegt vorzugsweise im Bereich von 4 bis 12 Nukleotiden, besonders bevorzugt von 5 bis 10 Nukleotiden. Der Stem der Stem-Loop-Struktur enthält vorzugsweise zwei vollständig zueinander komplementäre Abschnitte. Es können jedoch auch ein oder mehrere Basenfehlpaarungen vorhanden sein, sofern dadurch die 15 Stabilität nicht zu stark verringert wird. Die Basenpaarungen im Stem können AT- und GC-Basenpaarungen und Kombinationen davon sein. Ein Anteil der GC-Basenpaarungen von >50 % ist bevorzugt. Die Länge des Loops beträgt vorzugsweise 2 bis 8 Nukleotide, sie ist jedoch nicht sonderlich kritisch. Die thermodynamische Stabilität der Stem-Loop- 20 Struktur ist zweckmäßigerweise hoch genug, dass die Ausbildung einer Sekundärstruktur im Bereich des ATG-Startcodons, der 15 5'-liegenden Nukleotide, welche die Shine-Dalgarno-Sequenz beinhalten, und mindestens der 5 3'-befindlichen Nukleotide verhindern. Andererseits sollte die thermodynamische Stabilität der Stem-Loop-Struktur nicht so hoch sein, 25 dass die Prozessierung des Ribosoms auf der mRNA behindert wird. Vorzugsweise liegt die thermodynamische Stabilität der Stem-Loop-Struktur im Bereich von -4 bis -15 kcal/mol.

30 Die zur Expression des gewünschten Proteins verwendeten Expressionskontrollsequenzen umfassen Promotoren, ribosomale Bindungsstellen, d.h. Shine-Dalgarno-Sequenzen für prokaryontische Expressionssysteme bzw. Kozak-Sequenzen für eukaryontische

Expressionssysteme, Enhancer, Terminatoren, Polyadenylierungs-Sequenzen etc. Dem Fachmann sind entsprechende Expressionskontrollsequenzen aus Standardlehrbüchern der Molekularbiologie, z.B. Sambrook et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor oder Ausubel et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, bekannt.

Weiterhin kann die heterologe Nukleinsäuresequenz auch Abschnitte enthalten, die für eine Aufreinigungsdomäne, z.B. eine Poly-His-Domäne, eine FLAG-Epitop-Domäne etc., oder/und für eine Proteinase-Erkennungsdomäne, z.B. eine IgA-Protease- oder Faktor-X-Domäne, kodieren. Durch die Aufreinigungsdomäne kann die Gewinnung des gewünschten Proteins, z.B. aus einem *in vitro* Translationsansatz oder einer Wirtszelle bzw. dem zur Kultivierung verwendeten Medium, vereinfacht werden. Durch Proteinase-Spaltung innerhalb der Proteinase-Erkennungsdomäne kann die heterologe Peptidsequenz von dem gewünschten Protein abgespalten werden.

Die heterologe Nukleinsäuresequenz oder/und die für das gewünschte Protein kodierende Nukleinsäuresequenz werden - um eine weitere Verbesserung der Expressionshöhe zu erreichen - günstigerweise so gewählt, dass sie zumindest teilweise eine an das jeweilige Expressionssystem angepasste Codon-Nutzung aufweisen.

Ein weiterer Gegenstand der Erfindung ist ein Reagenz zur Herstellung eines Proteins, umfassend

(a) eine zu der für das gewünschte Protein kodierenden Nukleinsäuresequenz heterologe Nukleinsäuresequenz, die im korrekten Leseraster in die Protein-kodierende Nukleinsäuresequenz eingefügt werden kann, und die in einem Abstand von 6-30 Nukleotiden 3'-seitig des Translations-Startcodons eine Stem-Loop-Struktur ausbilden kann, und

(b) ein zur Herstellung des Proteins geeignetes Expressionssystem.

Die heterologe Nukleinsäuresequenz kann in Form einer vollständigen Sequenz oder in Form von mehreren Teilsequenzen vorliegen.

5

Das erfindungsgemäße Verfahren und Reagenz sind insbesondere für die Synthese von Proteinen schwer exprimierbarer Gene sowie der Synthese von Proteinen ausgehend von Genbanken anwendbar, da hierbei die Erfolgsrate gegenüber der bei der Verwendung üblicher Expressionsvektoren gesteigert werden kann.

10

Weiterhin soll die vorliegende Erfindung durch die nachfolgenden Abbildungen und Beispiele näher erläutert werden.

15

Es zeigen:

**Figur 1** eine schematische Darstellung der zur Durchführung einer Zweischritt-PCR notwendigen Nukleinsäure-Sequenzelemente;

20

**Figur 2** eine schematische Darstellung von unterschiedlich langen Stem-Loop-Strukturen in zur Einfügung in GFP Expressionkonstrukte verwendeten heterologen Nukleinsäure-Sequenzen;

25

**Figur 3** eine Auswertung der Ergebnisse bei der Expression von GFP unter Verwendung der Hairpin-Loop-GFP-Konstrukte gemäß Figur 3 in einem RTS-Expressionssystem. 1 µl jedes Ansatzes (Doppelbestimmungen) wurde elektrophoretisch über SDS-PAGE aufgetrennt und auf eine PVDF-Membran geblottet. Die Detektion erfolgte über DCP-Star und Lumi-Imager;

30

**Figur 4** eine schematische Darstellung von unterschiedlichen positionierten Stem-Loop-Strukturen in zur Einfügung von GFP-Expressionskonstrukte verwendeten heterologen Nukleinsäuresequenzen;

Figur 5 die Expression von GFP unter Verwendung der in Figur 4 dargestellten heterologen Nukleinsäure-Sequenzen. Die Durchführung und Auswertung der Ansätze erfolgte wie in der Legende zur Figur 3 beschrieben;

5

Figur 6 eine Auswertung der Ergebnisse bei der Expression des CIITA Gens (Wildtyp: Spur 1; Mutanten Spuren 2-10) unter Verwendung unterschiedlicher heterologer Nukleinsäuresequenzen mit Stem-Loop-Strukturen;

10

Figur 7 eine Auswertung der Ergebnisse bei der Expression des CMV Capsid (1049) Gens (Wildtyp: Spur 1; Mutanten Spuren 2-10) unter Verwendung unterschiedlicher heterologer Nukleinsäuresequenzen mit Stem-Loop-Strukturen;

15

Figur 8 eine Auswertung der Ergebnisse bei der Expression des Survivin Gens (Wildtyp: Spur 10; Mutanten Spuren 1-9) unter Verwendung unterschiedlicher heterologer Nukleinsäuresequenzen mit Stem-Loop-Strukturen;

20

Figur 9 eine Auswertung der Ergebnisse bei der Expression des GFP Gens (Wildtyp: Spur 10; Mutanten Spuren 1-9) unter Verwendung unterschiedlicher heterologer Nukleinsäuresequenzen mit Stem-Loop-Strukturen;

25

Figur 10 eine Auswertung der Ergebnisse bei der Expression des GFP und des 1049 Gens unter Verwendung unterschiedlicher heterologer Nukleinsäuresequenzen mit und ohne Stem-Loop-Strukturen;

30

Figur 11 eine Auswertung der Ergebnisse bei der Expression des CIITA und des Survivin Gens unter Verwendung unterschiedlicher heterologer Nukleinsäuresequenzen mit und ohne Stem-Loop-Strukturen;

**Figur 12 eine schematische Darstellung von zwei unterschiedlichen Stem-Loop-Strukturen in den erfindungsgemäßen heterologen Sequenzen;**

5      **Figur 13 eine Auswertung der Ergebnisse mit den in Figur 12 gezeigten Stem-Loop-Strukturen.**

### **Beispiele**

#### **Beispiel 1: Zweischritt-PCR**

10      Mit Hilfe einer Zweischritt-PCR können zu exprimierende Gene amplifiziert und mit den entsprechenden Kontrollregionen, wie T7-Promotor, T7 Gen10-Leader (g10), Ribosomaler Bindungsstelle (RBS) und T7-Terminator, versehen werden. Im ersten Schritt wird mittels eines Paares von Primern (A, B), die jeweils auf 15 Basen Länge mit dem entsprechenden Gen komplementär sind und 15 weitere Basen enthalten, die komplementär zu einem zweiten Primerpaar (C, D) sind, das Gen amplifiziert. Das zweite Primerpaar enthält alle wichtigen regulatorischen Elemente, die somit bei einer zweiten PCR-Amplifikation an das Gen angehängt werden (siehe Figur 1).

20      Über den A-Primer können bei dieser Methode Veränderungen im 5'-Bereich des Gens eingeführt werden. Bei den Hairpin-Loop-Konstrukten wurden über diesen A-Primer Hairpin-Loops mit unterschiedlicher Länge des Hairpin-Loop-Stamms und in unterschiedlicher Position hinter dem Start-Codon in die Gen-Sequenz inseriert.

#### **Primer C (SEQ ID NO. 1)**

##### T7-Promotor

30      5'-GAAATTAAATACGACTCACTATAGGGAGACCACAACGGTTCCCTCT

g10

RBS

AGAAATAATTTGTTAACCTTAAGAAGGAGATATACC-3'

**Komplementär zu A**

5

Primer D (SEQ ID NO. 2)

T7-Terminator

10 5'-CAAAAAACCCCTCAAGACCGTTAGAGGCCCAAGGGGGCCGC

CAGTGTGCTGAATTCGCCTTTATTA-3'

**Komplementär zu B**

15 Reaktionsbedingungen

Die PCR-Reaktionen wurden üblicherweise nach folgendem Schema mit dem Expand High Fidelity Kit (Roche Applied Science) im 50 µl Maßstab durchgeführt:

20 PCR 1: Template 10ng/Ansatz; Primer A 20 pmol/Ansatz; Primer B  
20 pmol/Ansatz

95 °C 5 min + 20 mal (95 °C 1 min + 55 1 min + 72 °C  
1 min) + 4 °C

PCR 2: 2 µl PCR 1; Primer C 20 pmol/Ansatz; Primer D 20  
pmol/Ansatz

95 °C 5 min + 30 mal (95 °C 1 min + 50 °C 1 min +  
72 °C 1 min) + 72 °C 10 min +4 °C

30 Beide PCR-Reaktionen wurden jeweils durch Agarose-Gel-Elektrophorese  
überprüft und die PCR-Produkte der PCR 2 gleichzeitig mit Hilfe eines  
DNA-Längenstandards, welcher definierte DNA-Mengen enthält, im Lumi-

Imager-System quantifiziert. Die daraus erhaltenen PCR-Produkte wurden direkt als Template in RTS-Expressions-Ansätzen eingesetzt.

**Beispiel 2: Expression mit dem RTS *in vitro* Expressionssystem**

5

Die Expressionen mit dem RTS 100 HY Kit (Firma Roche Applied Science) wurden in Batch-Ansätzen zu je 50 µl nach der dem Kit beiliegenden Anleitung durchgeführt. Dabei wurden DNA-Mengen von 0,25-1 µg pro Reaktions-Ansatz eingesetzt. Um die Ergebnisse einer Versuchsreihe vergleichen zu können, wurden immer gleiche Mengen des jeweiligen Templates eingesetzt. Die Ansätze wurden bei 30 °C für 4 h inkubiert.

10

**Beispiel 3: Expression von Hairpin-Loop-GFP-Konstrukten**

15

Es wurde am Beispiel von GFP (Grün-Fluoreszenz-Protein) untersucht, welchen Einfluss Hairpin-Loops (Haarnadel-förmige Schleifen) in der mRNA direkt nach dem Start-ATG ausüben. Dazu wurden RNA-Sequenzen ermittelt, die Hairpin-Loops (HL) mit unterschiedlichen Stammlängen ausbilden. Je länger der Stamm des Hairpin-Loops ist, um so energetisch stabiler ist diese Struktur. Bei der Erstellung der Hairpin-Loops wurde darauf geachtet, dass nur in *E. coli* Genen häufig zu findende Codons verwendet wurden. Die ermittelten Sequenzen für die verschiedenen Hairpin-Loops wurden mittels mRNA-Sekundärstrukturanalyse auf ihre Stabilität im gesamten Konstrukt hin überprüft. Für Sequenzen, die genügend Stabilität aufwiesen, wurden Primer erstellt, die in der beschriebenen Zweischritt-PCR gemäß Beispiel 1 eingesetzt wurden.

20

25

Primer A:

ohne Hairpin-Loop (SEQ ID NO. 3)

30

Komplementär zu C

GFP

5'-AGGAGATATACCATGACTAGCAAAGGAGAA-3'

Stammlänge 4 bp (SEQ ID NO.4)

Komplementär zu C      HL 4 bp                          GFP  
5'-AGGAGATATACCATGACTAATTTAGTACTAGCAAAGGAGAA-3'

5 Stammlänge 5 bp (SEQ ID NO. 5)

Komplementär zu C      HL 5 bp                          GFP  
5'-AGGAGATATACCATGACTGTTATACAGTAACTAGCAAAGGAGAA-3'

Stammlänge 6 bp (SEQ ID NO. 6)

Komplementär zu C      HL 6 bp                          GFP  
5'-AGGAGATATACCATGACTGGTCAATTACCAGTAACTAGCAAAGGAGAA-3'

Stammlänge 7 bp (SEQ ID NO. 7)

Komplementär zu C      HL 7 bp                          GFP  
5'-AGGAGATATACCATGACTGCTTACATCAAGCAGTAACTAGCAAAGGAGAA-3'

Stammlänge 8 bp (SEQ ID NO. 8)

Komplementär zu C      HL 8 bp                          GFP  
5'-AGGAGATATACCATGACTGCACGTGATCGTCAGTAACTAGCAAAGGAGAA-3'

Primer B (SEQ ID NO. 9)

Komplementär zu D      GFP  
5'-ATTCGCCCTTTATTAATGATGATGATGATG-3'

Eine schematische Darstellung der mRNA-Sekundärstrukturen der Hairpin-Loop-GFP-Konstrukte ist in Figur 2 gezeigt.

Nach Expression im RTS gemäß Beispiel 2 wurde zur Verifizierung die gebildete GFP-Menge im Fluorimeter gemessen und der Western-Blot mittels CDP-Star-Detektion und Auswertung im Lumi-Imager quantitativ analysiert. Die Ergebnisse sind in Figur 3 gezeigt.

5

Es ist eindeutig erkennbar, dass die Expressionsrate mit der Stammlänge des Hairpin-Loops variiert. Bis zu einer Stammlänge von 5 bp ist die Expressionsrate relativ konstant, um dann anschließend abzufallen. Bei einer Stammlänge von 8 bp ist nahezu keine Expression mehr nachweisbar.

10

Diese Untersuchungen bestätigen die oben erhaltenen Ergebnisse. Man kann also sagen, dass ein Hairpin-Loop mit einer Stammlänge ab 6 bp oder besser mit einer Freien Energie von -7,8 kcal/mol eine Struktur darstellt, die einen erheblichen Einfluss auf die Expression ausübt. Zu erklären ist dies damit, dass diese Struktur bei den Expressionsbedingungen stabil ist und somit das vorgelagerte Start-ATG nicht frei zugänglich ist.

15

#### **Beispiel 4: Ermittlung des minimalen Abstandes zum Start-ATG**

Um nun zu bestimmen, bis zu welchem Abstand ein solcher Hairpin-Loop einen Einfluss auf die Expression ausübt, wurde der Hairpin-Loop mit der Stammlänge 8 bp (Energie -11,8 kcal/mol) in Schritten zu je 3 Basen vom Start-ATG weg in die GFP-Sequenz verschoben. Die Sequenzen der auf diese Weise erhaltenen A-Primer war wie folgt:

25 Stammlänge 8 bp, 6 Basen in die GFP-Sequenz hineinverschoben (SEQ ID NO. 10):

HL 8 bp                  GFP

5'-AGG...ATGACTAGCACT...GTAAAAGGAGAAGAACCTT-3'

30 Stammlänge 8 bp, 9 Basen in die GFP-Sequenz hineinverschoben (SEQ ID NO. 11.):

HL 8bp                  GFP

5'-AGG...ATGACTAGCAAAACT...GTAGGAGAAGAACCTTTC-3'

Stammlänge 8 bp, 12 Basen in die GFP-Sequenz hineinverschoben (SEQ ID NO.12):

5

HL 8 bp

GFP

5'-AGG...ATGACTAGCAAAGGAAC...GTAGAAGAACCTTTCACT-3'

10

Stammlänge 8 bp, 15 Basen in die GFP-Sequenz hineinverschoben (SEQ ID NO. 13):

HL 8 bp

GFP

5'-AGG...ATGACTAGCAAAGGAGAAAC...GTAGAACTTTCACTGG-3'

15

Stammlänge 8 bp, 18 Basen in die GFP-Sequenz hineinverschoben (SEQ ID NO. 14):

HL 8 bp

GFP

5'-AGG...ATGACTAGCAAAGGAGAAGAAAC...GTACTTTCACTGGAG

20

TT-3'

Stammlänge 8 bp, 21 Basen in die GFP-Sequenz hineinverschoben (SEQ ID NO. 15):

HL 8bp

GFP

25

5'-AGG...ATGACTAGCAAAGGAGAAGAACCTTACT...GTATTCACTGGAG  
TTGTC-3'

30

Diese DNA-Konstrukte mit den in Figur 4 gezeigten Sekundärstrukturen wurden ebenfalls über Zweischritt-PCR mit den zuvor beschriebenen Primern B, C und D synthetisiert und direkt aus der PCR-Reaktion in Expressionsansätzen als Template eingesetzt. Dabei wurde durch Quantifizierung über ein Agarose-Gel mit DNA-Marker VII und Auswertung

dieses Gels im Lumi-Imager sichergestellt, dass gleiche Mengen an Template eingesetzt wurden. Die Expressionsansätze wurden über einen Western-Blot ausgewertet. Die Ergebnisse sind in Figur 5 gezeigt.

- 5 Die Expressionen lassen erkennen, dass ab einem Abstand von 9 Basen zum Start-ATG die Translation der mRNA möglich ist. Es besteht aber noch ein hemmender Einfluss des Hairpin-Loops. Erst ab einem Abstand von 12 Basen verläuft die Translation wieder nahezu ungehemmt. Man kann also aus diesen Ergebnissen schließen, dass das Ribosom einen Platzbedarf von  
10 9-11 Basen nach dem Start-ATG besitzt. Des Weiteren lässt sich aus diesen Ergebnissen folgern, dass ein Hairpin-Loop, der 12 oder mehr Basen nach dem Start-ATG entfernt ist, zwar einen Einfluss auf die mRNA-Sekundärstruktur aber keinen auf die Initiation der Expression hat.

15 **Beispiel 5: Einführung von Stem-Loop-Strukturen zur Auflösung ungünstiger Sekundärstrukturen**

Bei früheren Expressionsansätzen mit dem Rapid Translation System (Roche Applied Science) wurde bei einigen Genen nur eine geringe oder gar 20 keine Expression gefunden. Als Ursache hierfür wurde oft eine ungünstige RNA-Sekundärstruktur ermittelt, bei der entweder das Startcodon, oder die Shine-Dalgarno-Sequenz in einer Sekundärstruktur mit der Gensequenz involviert war, und damit in gebundener Form vorlag.

- 25 Für drei dieser Gene, Survivin, Cytomegalovirus Capsid Protein 1049 (1049) und Class II Transactivator (CIITA) wurde eine heterologe Nukleinsäuresequenz mit einem Hairpin-Loop und einer Stammlänge von 7 Basen in einem Abstand von 15 Basen nach dem Startcodon eingeführt. Direkt anschließend an den Hairpin Loop wurde das Wildtyp-Gen (siehe 30 unten \*) ohne das Start ATG gesetzt. Vor den Hairpin- Loop wurden jeweils AT reiche Sequenzen gesetzt, welche weniger stabile Basenpaarungen als GC reiche Sequenzen ausbilden können. Außerdem

wurde darauf geachtet keine für E. coli seltenen Codons innerhalb der eingeführten Sequenzen zu verwenden.

Dadurch, dass nun einerseits ein stabiler idealer Hairpin-Loop vorgegeben wird und andererseits direkt nach dem Startcodon eine Sequenz folgt, welche nicht zu Sekundärstrukturbildung neigt, sollte unabhängig vom nachfolgenden Gen eine freie Zugängigkeit für den Initiationskomplex mit der kleinen ribosomale Untereinheit an die Shine-Dalgarno-Sequenz und das Start-ATG ermöglicht werden.

10

Es wurden jeweils 9 verschiedene AT-reiche Sequenzen vor den Hairpin-Loops eingesetzt und mit den Wildtyp-Genen\* verglichen. Als Kontrollgen wurde das GFP cycle 3 Protein mit den gleichen Hairpin-Loops und AT-reichen Sequenzen über die in Beispiel 1 erwähnte Zweischritt-PCR synthetisiert. Die Sequenzen der A und B Primer sind unten angegeben. In Primer 1 wurden jeweils die homologen Bereiche zu Primer C unterstrichen. Die AT-reiche Sequenz wurde in kursiv, der Hairpin-Loop in Fettdruck und die Wildtyp Gensequenz in fett und unterstrichen angegeben. In Primer B wurden die zu Primer D homologen Bereiche unterstrichen und die Bereiche mit Homologie zum Wildtyp Gen fett angegeben. Als Primer D wurde abweichend zu Beispiel 1 der folgende Primer benutzt:

Primer D (SEQ ID NO 16):

CAAAAAACCCCTCAAGACCCGTTAGAGGGCCCCAAGGGGTTGGGAGTA

25 GAATGTTAAGGATTAGTTTATTA

Die unterstrichene Region ist homolog zu Primer C.

Varianten Primer A:

1049 - 1 (SEQ ID NO. 17):

30 AGGAGATATACCATGAAATATACATATTCTCTGCACGTGATCGTGCAG  
GCTAACACCGCG

**1049 - 2 (SEQ ID NO: 18):**

AGGAGATATACCATGAAAACATATTATTCTCTGCACGTGATCGTGCAG  
GCTAACACCGCG

5      **1049 - 3 (SEQ ID NO. 19):**

AGGAGATATACCATGAAATATTCTTATAACTGCACGTGATCGTGCAG  
GCTAACACCGCG

10

**1049 - 4 (SEQ ID NO. 20):**

AGGAGATATACCATGAAATATTATTCTACACTGCACGTGATCGTGCAG  
GCTAACACCGCG

15

**1049 - 5 (SEQ ID NO. 21):**

AGGAGATATACCATGAAATATACATATTCACTGCACGTGATCGTGCAG  
GCTAACACCGCG

20

**1049 - 6 (SEQ ID NO. 22):**

AGGAGATATACCATGAAAACATATTATTCACTGCACGTGATCGTGCAG

GCTAACACCGCG

**1049 - 7 (SEQ ID NO. 23):**

AGGAGATATACCATGAAATATTCAATACACTGCACGTGATCGTGCAG

25

GCTAACACCGCG

**1049 - 8 (SEQ ID NO. 24):**

AGGAGATATACCATGAAATATTATTCAACACTGCACGTGATCGTGCAG

GCTAACACCGCG

30

**1049 - 9 (SEQ ID NO. 25):**

AGGAGATATACCATGCATCATCATCATCATCTGCACGTGATCGTGCAG  
GCTAACACCGCG

**1049 – 10 (Wildtyp) (SEQ ID NO. 26):**

5 **AGGAGATATACCATGGCTAACACCGCG**

**1049 – Primer B (SEQ ID NO. 27):**

**AGGATTAGTTTATTAATATGATGATGATGATGGCGCCGGTGCGCGA**

**Unterstrichenes ist homolog zu Primer D**

10

**Varianten Primer A:**

**Survivin - 1 (SEQ ID NO. 28):**

AGGAGATATACCATGAAATATACATATTCTCTGCACGTGATCGTGCAG

15

**GGTGCCCCGACG**

**Survivin - 2 (SEQ ID NO. 29):**

AGGAGATATACCATGAAAACATATTATTCTCTGCACGTGATCGTGCAG

**GGTGCCCCGACG**

20

**Survivin - 3 (SEQ ID NO. 30):**

AGGAGATATACCATGAAATATTCTTATAACTGCACGTGATCGTGCAG

**GGTGCCCCGACG**

25

**Survivin - 4 (SEQ ID NO. 31):**

AGGAGATATACCATGAAATATTATTCTACACTGCACGTGATCGTGCAG

**GGTGCCCCGACG**

**Survivin - 5 (SEQ ID NO. 32):**

30

AGGAGATATACCATGAAATATACATATTCACTGCACGTGATCGTGCAG

**GGTGCCCCGACG**

**Survivin - 6 (SEQ ID NO. 33):**

AGGAGATATACCATGAAAACATATTATTCACTGCACGTGATCGTGCAG  
GGTGCCCCGACG

5      **Survivin - 7 (SEQ ID NO. 34):**

AGGAGATATACCATGAAATATTCAATACACTGCACGTGATCGTGCAG  
GGTGCCCCGACG

**Survivin - 8 (SEQ ID NO. 35):**

10     AGGAGATATACCATGAAATATTCAACACTGCACGTGATCGTGCAG  
GGTGCCCCGACG

**Survivin - 9 (SEQ ID NO. 36):**

15     AGGAGATATACCATGCATCATCATCATCATCTGCACGTGATCGTGCAG  
GGTGCCCCGACG

**Survivin – 10 (A Wildtyp) (SEQ ID NO. 37):**

AGGAGATATACCATGGGTGCCCGACG

20     **Survivin – Primer B (SEQ ID NO. 38):**

AGGATTAGTTATTAATGATGATGATGATGATGATCCATGGCAGCCAGC

**CIITA - 1 (SEQ ID NO: 39):**

AGGAGATATACCATGAAAATACATATTCTCTGCACGTGATCGTGCAG  
25     GAGTTGGGGCCC

**CIITA - 2 (SEQ ID NO. 40):**

AGGAGATATACCATGAAAACATATTATTCTCTGCACGTGATCGTGCAG  
GAGTTGGGGCCC

30

**CIITA - 3 (SEQ ID NO. 41):**

AGGAGATATACCATGAAATATTCTTATAACTGCACGTGATCGTCAG  
GAGTTGGGGCCC

**CIITA - 4 (SEQ ID NO: 42):**

5 AGGAGATATACCATGAAATATTATTACTACACTGCACGTGATCGTCAG  
GAGTTGGGGCCC

**CIITA - 5 (SEQ ID NO. 43):**

10 AGGAGATATACCATGAAATATACTACATATTCACTGCACGTGATCGTCAG  
GAGTTGGGGCCC

**CIITA - 6 (SEQ ID NO. 44):**

15 AGGAGATATACCATGAAAACATATTATTCACTGCACGTGATCGTCAG  
GAGTTGGGGCCC

**CIITA - 7 (SEQ ID NO. 45):**

20 AGGAGATATACCATGAAATATTCTATACACTGCACGTGATCGTCAG  
GAGTTGGGGCCC

**CIITA - 8 (SEQ ID NO. 46):**

25 AGGAGATATACCATGAAATATTATTCAACACTGCACGTGATCGTCAG  
GAGTTGGGGCCC

**CIITA - 9 SEQ ID NO. 47):**

AGGAGATATACCATGCATCATCATCATCATCTGCACGTGATCGTCAG  
GAGTTGGGGCCC

30 **CIITA - 10 (A Wildtyp) (SEQ ID NO: 48):**  
**AGGAGATATACCATGGAGTTGGGGCCC**

**CIITA Primer B (SEQ ID NO. 49):**

**AGGATTAGTTTATTATTAATGATGATGATGATGATGAGAACCCCC**

Im Folgenden sind die Sequenzen der über PCR erzeugten Expressionskonstrukte für die Mutante 1 und den Wildtyp angegeben. Die Wildtyp-Gensequenz ist fett angegeben. Am Ende des Gens wurde mit dem B-Primer jeweils ein hexa-Histidin-Tag zur Detektion mit einem spezifischen Antikörper eingeführt (unterstrichen).

**1049 – 1 (431 bp) (SEQ ID NO. 50):**

GAAATTAAATACGACTCACTATAAGGGAGACCACAACGGTTCCCTCTAG  
AAATAATTGTAACTTAAGAAGGAGATATACCATGAAATATACAT  
ATTCTCTGCACGTGATCGTCAGGCTAACACCGCGCCGGGACCCAGGG  
TGGCAAACAAGCGGGACGAAAAACACCGTCACGTCGTTAACGTCGTT  
TCCAGCTGCCGACCGAGATATCAGAGGCCACCCACCCGGTGTGGACA  
CCATGCTGAGCAAGTACACGCGCATGTCCAGCCTGTTAATGACAAGT  
GCGCCTTAAGCTGGACCTGTTGCGCATGGTAGCCGTGTCGCGCACCC  
**GGCGCCATCATCATCATCATCATCATCATCATCATCATCATCATCAT**  
TCCAACCCCTGGGCCTCTAAACGGGTCTTGAGGGTTTTG

**1049 – 10 (WildTyp) (398 bp) (SEQ ID NO. 51):**

GAAATTAAATACGACTCACTATAAGGGAGACCACAACGGTTCCCTCTAG  
AAATAATTGTAACTTAAGAAGGAGATATACCATGGCTAACACCG  
CGCCGGGACCCACGGTGGCCAACAAGCGGGACGAAAAACACCGTCAC  
GTCGTTAACGTCGTTGGAGCTGCCACCGAGATATCAGAGGCCACC  
CACCCGGTGTGGCCACCATGCTGAGCAAGTACACGCGCATGTCCAGC  
CTGTTAATGACAAGTGCCTTAAGCTGGACCTGTTGCGCATGGTAG  
CCGTGTCGCGCACCCGGCGCCATCATCATCATCATCATCATCATCAT  
ATCCTAACATTCTACTCCAACCCCTGGGCCTCTAAACGGGTCTTG  
AGGGGTTTTG

Survivin – 1 (632 bp) (SEQ ID NO. 52):

GAAATTAATACGACTCACTATAAGGGAGACCACAACGGTTCCCTCTAG  
AAATAATTGTTAACCTTAAGAAGGAGATATACCATGAAATATACAT  
ATTCTCTGCACGTGATCGTGAGGGTCCCCGACGTTGCCCTGCCTG  
5 GCAGCCCTTCTCAAGGACCACCGCATCTCTACATTCAAGAACTGGCC  
CTTCTGGAGGGCTGCCTGCACCCGGAGCGGATGCCGAGGCTGG  
CTTCATCCACTGCCCACTGAGAACGAGCCAGACTTGGCCAGTGT  
CTTCTGCTTCAAGGAGCTGGAAGGCTGGAGCCAGATGACGACCCAT  
AGAGGAACATAAAAAGCATTGTCGGTTGCGCTTCCTTCTGTCAAG  
10 AAGCAGTTGAAGAATTAACCCTGGTGAATTGGAAACTGGACAGA  
GAAAGAGCCAAGAACAAAATTGCAAAGGAAACCAACAATAAGAAGAA  
AGAATTGAGGAAACTGCGAAGAAAGTGCCTGCCATCGAGCAGCT  
GGCTGCCATGGATCATCATCATCATCATATTAAATAACTAATCCTTAA  
CATTCTACTCCCAACCCCTGGGCCTCTAAACGGGTCTTGAGGGGTTT  
15 TTTG

Survivin – 10 (Wildtyp) (599 bp) (SEQ ID NO. 53):

GAAATTAATACGACTCACTATAAGGGAGACCACAACGGTTCCCTCTAG  
AAATAATTGTTAACCTTAAGAAGGAGATATACCATGGGTGCCCGA  
20 CGTTGCCCTGCCTGGCAGCCCTTCTCAAGGACCACCGCATCTCTAC  
ATTCAAGAACTGCCCTTCTGGAGGGCTGCCTGCACCCGGAGCG  
GATGCCGAGGCTGGCTTCATCCACTGCCCACTGAGAACGAGCCAGA  
CTTGGCCCAGTGTCTTGCTTCAAGGAGCTGGAAGGCTGGAGCC  
AGATGACGACCCATAGAGGAACATAAAAAGCATTGTCGGTTGCGC  
25 TTTCTTCTGTCAAGAAGCAGTTGAAGAATTAACCCTGGTGAATT  
TTGAAACTGGACAGAGAAAGAGCCAAGAACAAAATTGCAAAGGAAAC  
CAACAATAAGAAGAAAGAATTGAGGAAACTGCGAAGAAAGTGCCTC  
GTGCCATCGAGCAGCTGGCTGCCATGGATCATCATCATCATCATATT  
ATAAAACTAATCCTTAACATTCTACTCCCAACCCCTGGGCCTCTAAAC  
30 GGGTCTTGAGGGGTTTTG

CIITA – 1 (1400 bp) (SEQ ID NO. 54):

GAAATTAAATACGACTCACTATAAGGGAGACCACAACGGTTCCCTCTAG  
AAATAATTGTAACTTAAGAAGGAGATATACCATGAAATATA  
ATTCTCTGCACGTGATCGTCAGGAGTTGGGGCCCTAGAAGGTGGCT  
ACCTGGAGCTTCTAACAGCGATGCTGACCCCCCTGTGCCCTACCACTT  
5 CTATGACCAGATGGACCTGGCTGGAGAAGAAGAGATTGAGCTACTC  
AGAACCCGACACAGACACCATCAACTGCGACCAGTCAGCAGGCTGTT  
GTGTGACATGGAAGGTGATGAAGAGACCAGGGAGGCTTATGCCAATAT  
CGCGGAACCTGGACCAGTATGTCTCCAGGACTCCCAGCTGGAGGGCCT  
GAGCAAGGACATTTCAAGCACATAGGACCAGATGAAGTGATCGGTGA  
10 GAGTATGGAGATGCCAGCAGAAGTTGGCAGAAAAGTCAGAAAAGAC  
CCTTCCCAGAGGAGCTTCCGGCAGACCTGAAGCACTGGAAGCCAGCTG  
AGCCCCCCCAGTGTGGTACTGGCAGTCTCCTAGTGGACCAGTGAGCG  
ACTGCTCCACCCCTGCCCTGCCACTGCCTGCGCTGTTCAACCAGG  
AGCCAGCCTCCGGCCAGATGCCCTGGAGAAAACCGACCAGATTCCC  
15 ATGCCTTCTCCAGTTCTCGTTGAGCTGCCTGAATCTCCCTGAGGGAC  
CCATCCAGTTGTCCCCACCATCTCCACTCTGCCCATGGCTCTGGCA  
AATCTCTGAGGCTGGAACAGGGGTCTCCAGTATATTCTACCATGGT  
GAGGTGCCCAAGGCCAGCAAGTACCCCTCCAGTGGATTCACTGTC  
CACGGCCTCCAACATCTCCAGACCGGCCAGGCTCCACCAGCCCCCTC  
20 GCTCCATGCCACTGACCTGCCAGCATGCCGAACCTGCCCTGACC  
TCCCGAGCAAACATGACAGAGCACAAGACGTCCCCACCAATGCC  
GGCAGCTGGAGAGGTCTCCAACAAGCTTCAAAATGGCCTGAGCCGGT  
GGAGCAGTTCTACCGCTCACTGCAGGACACGTATGGTGCCGAGCCC  
AGGCCCGGATGGCATCCTAGTGGAGGTGGATCTGGTGCAGGCCAGGCT  
25 GGAGAGGAGCAGCAGCAAGAGCCTGGAGCGGGAACTGCCACCCGG  
ACTGGGCAGAACGGCAGCTGGCCAAGGAGGCCGGCTGGCTGAGGTGCTG  
TTGGCTGCCAAGGAGCACCGGCCGGCGTCGACTCGAGCGAGCTCCC  
GGGGGGGGTTCTCATCATCATCATCATTTAATAATAACTAATCCTT  
AACATTCTACTCCCAACCCCTGGGCCTCTAAACGGGTCTTGAGGGGG  
30 TTTTTTG

CIITA -10 WT 1367 bp (SEQ ID NO. 55):

GAAATTAAATACGACTCACTATAAGGGAGACCACAACGGTTCCCTCTAG  
AAATAATTTGTTAACCTTAAGAAGGAGATATACCATGGAGTTGGGC  
CCCTAGAAGGTGGCTACCTGGAGCTTCTAACAGCGATGCTGACCCCC  
TGTGCCTCTACCACTTCTATGACCAGATGGACCTGGCTGGAGAAGAAG  
5 AGATTGAGCTCTACTCAGAACCCGACACAGACACCATCAACTGCGACC  
AGTCAGCAGGCTGTTGTGACATGGAAGGTGATGAAGAGACCAGGG  
AGGCTTATGCCAATATCGCGGAACTGGACCAGTTGCTTCCAGGACT  
CCCAGCTGGAGGGCCTGAGCAAGGACATTTCAAGCACATAGGACCAG  
ATGAAGTGATCGGTGAGAGTATGGAGATGCCAGCAGAAGTTGGCAG  
10 AAAAGTCAGAAAAGACCCTCCAGAGGGAGCTTCCGGCAGACCTGAA  
GCACTGGAAGCCAGCTGAGCCCCACTGTGGTGACTGGCAGTCTCCT  
AGTGGGACCAGTGAGCGACTGCTCCACCCCTGCCCTGCCACTGCC  
TGCCTGTTCAACCAGGAGCCAGCCTCCGGCCAGATGCCCTGGAGAA  
AACCGACCAGATTCCCATGCCTTCTCCAGTTCTCGTTGAGCTGCC  
15 AATCTCCCTGAGGGACCCATCCAGTTGTCCCCACCATCTCCACTCTGC  
CCCATGGGCTCTGGCAAATCTCTGAGGCTGGAACAGGGTCTCCAGTA  
TATTCATCTACCATGGTGAGGTGCCAGGCCAGCCAAGTACCCCTC  
CCAGTGGATTCACTGTCCACGGCCTCCAAACATCTCCAGACGGCCAG  
GCTCCACCAGCCCCCTCGCTCCATGCCACTGACCTGCCAGCATGC  
20 CTGAACCTGCCCTGACCTCCGAGCAAACATGACAGAGCACAAGACGT  
CCCCCACCAATGCCGGCAGCTGGAGAGGTCTCAACAAAGCTTCAA  
AATGGCCTGAGCCGGTGGAGCAGTTCTACCGCTCAGTGAGGACACGT  
ATGGTGCCGAGCCGCAGGCCGGATGGCATCCTAGTGGAGGTGGATC  
TGGTGAGGCCAGGCTGGAGAGGGAGCAGCAGCAAGAGCCTGGAGCGG  
25 GAACTGGCCACCCGGACTGGCAGAACGGCAGCTGCCCAAGGAGG  
CCTGGCTGAGGTGCTGTTGGCTGCCAAGGAGCACCGGGCCCGTCG  
ACTCGAGCGAGCTCCGGGGGGTTCTCATCATCATCATCATCAT  
ATAATAAAACTAACCTAACATTCTACTCCAAACCCCTGGGGCCTCTA  
AACGGGTCTTGAGGGTTTTTG

30

GFP CyC3 – 1 (938 bp) (SEQ ID NO. 56):

GAAATTAAATACGACTCACTATAAGGGAGACCACAACGGTTCCCTCTAG  
AAATAATTTGTTAACTTAAGAAGGAGATATACCATGAAATATACAT  
ATTCTCTGCACGTGATCGTCAGACTAGCAAAGGAGAAGAACTTTCA  
CTGGAGTTGTCCCAATTCTTGTGAATTAGATGGTATGTTAATGGGCA  
5 CAAATTTCTGTCAGTGGAGAGGGTGAAGGTGATGCTACATACGGAAA  
GCTTACCCCTAAATTTATTGCACTACTGGAAAACACTACCTGTTCCATGG  
CCAACACTTGTCACTACTTCTTATGGTGTCAATGCTTTCCCGTTA  
TCCGGATCATATGAAACGGCATGACTTTCAAGAGTGCCATGCCGA  
AGGTTATGTACAGGAACGCACATATCTTCAAAGATGACGGGAACTA  
10 CAAGACCGGTGCTGAAGTCAAGTTGAAGGTGATACCCCTGTTAATCG  
TATCGAGTTAAAAGGTATTGATTTAAAGAAGATGGAAACATTCTCGG  
ACACAAACTCGAGTACAACACTAACTCACACAATGTATACATCACGGC  
AGACAAACAAAAGAATGGAATCAAAGCTAACTCAAATTCGCCACA  
ACATTGAAGATGGATCCGTTCAACTAGCAGACCATTATCAACAAAATA  
15 CTCCAATTGGCGATGCCCTGTCCTTACCAAGACAACCATTACCTGTC  
GACACAATCTGCCCTTCGAAAGATCCAACGAAAAGAGAGACCACAT  
GGTCCTTCTTGAGTTGTAACAGCTGCTGGGATTACACATGGCATGGAT  
GAACATACAAACCCGGGGGGGTTCTCATCATCATCATCAT  
TAAACTAACCTTAACATTCTACTCCAAACCCCTGGGCCTCTAAACG  
GGTCTTGAGGGGTTTTTG  
20

**pIVEX-GFP CyC3 – 10 905 bp (SEQ ID NO. 57):**

GAAATTAAATACGACTCACTATAAGGGAGACCACAACGGTTCCCTCTAG  
AAATAATTTGTTAACTTAAGAAGGAGATATACCATGACTAGCAA  
25 GGAGAAGAACCTTCACTGGAGTTGCCAATTCTTGTGAATTAGATG  
GTGATGTTAATGGGCACAAATTTCTGTCAGTGGAGAGGGTGAAGGTG  
ATGCTACATACGGAAAGCTTACCCCTAAATTATTCGACTACTGGAAA  
ACTACCTGTTCCATGCCAACACTGTCACTAACCTCTTATGGTGTTC  
AATGCTTTCCCGTTATCCGGATCATATGAAACGGCATGACTTTCAA  
30 GAGTGCCATGCCGAAGGTTATGTACAGGAACGCACATATCTTCAA  
AGATGACGGGAACACTACAAGACGCGTGCTGAAGTCAAGTTGAAGGTGA  
TACCCCTGTTAACCGTATCGAGTAAAGGTATTGATTTAAAGAAGAT

GGAAACATTCTCGGACACAAACTCGAGTACAACATAACTCACACAAT  
GTATACATCACGGCAGACAAACAAAAGAATGGAATCAAAGCTAACTTC  
AAAATTGCCACAACATTGAAGATGGATCCGTTCAACTAGCAGACCAT  
TATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTACCAAGACA  
5 ACCATTACCTGTCGACACAATCTGCCCTTCGAAAGATCCAACGAAA  
AGAGAGACCACATGGTCCTTCTTGAGTTGTAACAGCTGCTGGGATTAC  
ACATGGCATGGATGAACTATAACAAACCCGGGGGGTTCATCATCA  
TCATCATCATTAATAAAACTAACCTAACATTACTCCCACCCCTTG  
GGGCCTCTAACGGGTCTTGAGGGGTTTTG

10

Man kann an den in den Figuren 6 bis 9 gezeigten Expressionen erkennen,  
dass die mit den Stem-Loop-Strukturen synthetisierten DNA Templates in  
allen Fällen zu einer Proteinsynthese führten, während beim Wildtypgen  
keine Proteinsynthese stattfand. Die Mutante 9 mit der Hexa-Histidin  
15 Sequenz wird dabei etwas schlechter als die übrigen AT-reichen Sequenzen  
exprimiert, hat aber den Vorteil, dass man das entstehende Protein über  
diese Markierung mit sechs Histidinresten an Ni-NTA-Chelat Säulen  
aufreinigen kann. Auch im Falle des GFP-Gens, einem ohnehin gut  
exprimierten Gen, führten die Stem-Loop-Konstrukte zu einer  
20 Ausbeutesteigerung.

20

#### Beispiel 6: Entfernen der Stem-Loop-Struktur zum Nachweis der Funktion

25

Um die Wirkung der Stem-Loop-Struktur von der Wirkung der eingeführten  
AT-reichen Sequenz zu unterscheiden wurden von je zwei der Mutanten  
eine identische PCR jedoch ohne den Teil des Stem-Loops hergestellt und  
im direkten Vergleich mit den Stem-Loop-Mutanten exprimiert.

30

An diesen Beispielen lässt sich deutlich der Einfluss der Stem-Loop-Struktur  
ersehen. Während beim GFP die AT-reiche Sequenz alleine zu einer  
gesteigerten Expression führt, bringt die Stem-Loop-Sequenz bei den  
schwer exprimierbaren Genen den entscheidenden Beitrag.

**Beispiel 7: Modifikation der Stem-Loop-Struktur zur Ermittlung der für die Funktion wichtigen Eigenschaften**

Um die Wirkung der GC-Basen innerhalb der Stem-Loop-Struktur zu  
5 ermitteln, wurde deren Sequenz durch eine AT-reiche Sequenz mit der gleichen freien Energie wie der des GC-reichen Stem-Loops ersetzt.

Für die Beispiele Survivin, CIITA und 1049 wurden dafür statt der ursprünglichen Stem-Loop-Sequenz CTG.CAC.GTG.ATC.GTG.CAG mit (G = -9,8 kcal/mol und einer Stammlänge von 7 Basenpaaren ein neuer Stem-  
10 Loop (Loop') mit der Sequenz CAG.ACA.AAT.AGA.TAT.TTG.TCT.GTA mit (G = -9,8 kcal/mol und einer Stammlänge von 9 Basenpaaren mit der AT-reichen Sequenz von Mutante 1 kombiniert. Die beiden Strukturen sind in Figur 12 gezeigt.

15

Man kann erkennen, dass beide Stem-Loop-Varianten die Expression gegenüber den jeweiligen Wildtypgenen deutlich steigern oder erst ermöglichen. Die GC-reichen Stem-Loop-Varianten weisen dabei eine etwas deutlichere Expressionssteigerung auf.

20

### **Ansprüche**

- 1. Verfahren zur Herstellung eines Proteins, umfassend die Schritte:**
  - (a) Bereitstellen einer für das Protein kodierenden Nukleinsäuresequenz, wobei 3'-seitig des Translations-Startcodons eine heterologe Nukleinsäuresequenz im korrekten Leseraster eingefügt wird, die so gewählt wird, dass in einem Abstand von 6-30 Nukleotiden 3'-seitig des Translations-Startcodons eine Stem-Loop-Struktur ausgebildet wird,**
  - (b) Bereitstellen eines zur Expression des Proteins geeigneten Expressionssystems und**
  - (c) Einbringen der Nukleinsäuresequenz gemäß (a) in das Expressionssystem gemäß (b) unter Bedingungen, dass das Protein synthetisiert wird.**
- 2. Verfahren nach Anspruch 1, weiterhin umfassend das Gewinnen des Proteins.**
- 3. Verfahren nach Anspruch 1 oder 2,  
dadurch gekennzeichnet,  
dass die eingefügte heterologe Nukleinsäuresequenz eine Länge bis zu 201 Nukleotiden aufweist.**
- 4. Verfahren nach Anspruch 3,  
dadurch gekennzeichnet,  
dass die eingefügte heterologe Nukleinsäuresequenz eine Länge bis zu 45 Nukleotiden aufweist.**

5. Verfahren nach einem der Ansprüche 1 bis 4,  
dadurch gekennzeichnet,  
dass die Stem-Loop-Struktur in einem Abstand von 12-21  
Nukleotiden 3'-seitig des Startcodons ausgebildet wird.

5

6. Verfahren nach einem der Ansprüche 1 bis 5,  
dadurch gekennzeichnet,  
dass die Länge des Stem in der Stem-Loop-Struktur im Bereich von  
4-12 Nukleotiden ist.

10

7. Verfahren nach einem der Ansprüche 1 bis 6,  
dadurch gekennzeichnet,  
dass der Bereich der heterologen Nukleinsäuresequenz, der 5'-seitig  
der Stem-Loop-Struktur liegt, selbst keine Sekundärstruktur ausbildet  
15 und keine Sekundärstruktur mit der 5'-untranslatierten Region der für  
das herzustellende Protein kodierenden Nukleinsäuresequenz  
eingehen kann.

15

8. Verfahren nach einem der Ansprüche 1 bis 7,  
dadurch gekennzeichnet,  
dass der Bereich der heterologen Nukleinsäuresequenz, der 5'-seitig  
der Stem-Loop-Struktur und 3'-seitig des ATG-Startcodons liegt,  
einen GC-Gehalt von < 50 % aufweist.

20

25 9. Verfahren nach einem der Ansprüche 1 bis 8,  
dadurch gekennzeichnet,  
dass man ein *in vitro* Expressionssystem verwendet.

25

10. Verfahren nach Anspruch 9,  
dadurch gekennzeichnet,  
dass man ein prokaryontisches *in vitro* Expressionssystem  
verwendet.

30

11. Verfahren nach Anspruch 10,  
dadurch gekennzeichnet,  
dass das prokaryontische *in vitro* Expressionssystem Lysate von  
gram-negativen Bakterien, insbesondere von *Escherichia coli*, oder  
gram-positiven Bakterien, insbesondere *Bacillus subtilis*, enthalten.  
5
12. Verfahren nach Anspruch 9,  
dadurch gekennzeichnet,  
dass man ein eukaryontisches *in vitro* Expressionssystem  
10 verwendet.
13. Verfahren nach Anspruch 12,  
dadurch gekennzeichnet,  
dass das eukaryontische *in vitro* Expressionssystem Lysate von  
Säugerzellen, insbesondere von Kaninchen, Reticulocyten, humanen  
15 Tumorzelllinien, Hamsterzelllinien, oder anderen Wirbeltierzellen,  
insbesondere Oozyten und Eiern von Fischen und Amphibien, sowie  
Insektenzelllinien, Hefezellen, Algenzellen oder Extrakte aus  
Pflanzenkeimen enthält.  
20
14. Verfahren nach einem der Ansprüche 1 bis 8,  
dadurch gekennzeichnet,  
dass man ein prokaryontisches *in vivo* Expressionssystem  
25 verwendet.
15. Verfahren nach Anspruch 14,  
dadurch gekennzeichnet,  
dass man eine prokaryontische Wirtszelle als Expressionssystem  
verwendet.  
30

16. Verfahren nach Anspruch 15,  
dadurch gekennzeichnet,  
dass man eine gram-negative prokaryontische Wirtszelle,  
insbesondere eine E.coli Zelle, oder eine gram-positive  
5 prokaryontische Wirtszelle, insbesondere eine *Bacillus subtilis* Zelle,  
verwendet.
  
17. Verfahren nach Anspruch 14,  
dadurch gekennzeichnet,  
10 dass man eine eukaryontische Wirtszelle als Expressionssystem  
verwendet.
  
18. Verfahren nach Anspruch 17,  
dadurch gekennzeichnet,  
15 dass man eine Hefezelle, eine Insektenzelle oder eine Wirbeltierzelle,  
insbesondere eine Amphibien-, Fische-, Vogel- oder Sägerzelle,  
verwendet.
  
19. Verfahren nach Anspruch 14,  
dadurch gekennzeichnet,  
20 dass man einen nicht-humanen eukaryontischen Wirtsorganismus als  
Expressionssystem verwendet.
  
20. Verfahren nach einem der Ansprüche 1 bis 19,  
dadurch gekennzeichnet,  
25 dass das Bereitstellen der für das Protein kodierenden  
Nukleinsäuresequenz durch Klonierung, Rekombination oder/und  
Amplifikation erfolgt.
  
- 30 21. Verfahren nach Anspruch 20,  
dadurch gekennzeichnet,  
dass das Bereitstellen eine Zweistufen-PCR umfasst.

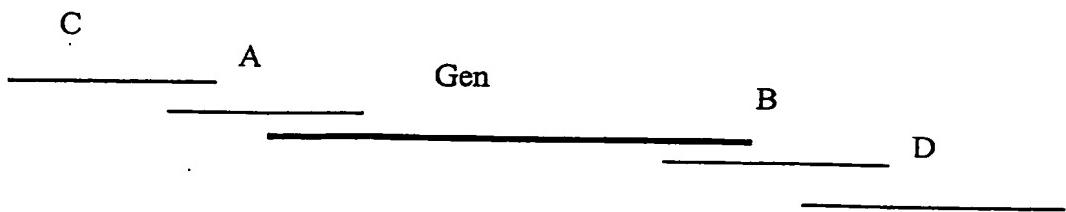
22. Verfahren nach einem der Ansprüche 1 bis 21,  
dadurch gekennzeichnet,  
dass die für das herzustellende Protein kodierende  
5 Nukleinsäuresequenz oder/und die heterologe Nukleinsäuresequenz  
zumindest teilweise eine an das jeweiligen Expressionssystem  
angepasste Codon-Nutzung aufweisen.
23. Verfahren nach einem der Ansprüche 1 bis 22,  
dadurch gekennzeichnet,  
10 dass die heterologe Nukleinsäuresequenz einen für eine  
Aufreinigungsdomäne oder/und einen für eine Proteinase-  
Erkennungsdomäne kodierenden Abschnitt enthält.
24. Reagenz zur Herstellung eines Proteins, umfassend  
15 (a) eine zu der für das Protein kodierenden Nukleinsäuresequenz  
heterologe Nukleinsäuresequenz, die im korrekten Leseraster  
in die Protein-kodierende Nukleinsäuresequenz eingefügt  
werden kann, und die in einem Abstand von 6-30 Nukleotiden  
3'-seitig des Translations-Startcodons eine Stem-Loop-  
20 Struktur ausbilden kann, und  
(b) ein zur Herstellung des Proteins geeignetes  
Expressionssystem.

### **Zusammenfassung**

Die Erfindung betrifft ein Verfahren zur optimierten Herstellung von  
5 Proteinen in einem *in vitro* oder *in vivo* Expressionssystem sowie dafür  
geeignete Reagenzien.

10 pu/ANM/29415P DE-09.12.2002

**Figur 1**



**Figur 2**

HL mit Stammlänge 6bp: -7,8 kcal/mol

A  
A U  
C U  
U-A  
G-C  
G-C  
U-A  
C-G  
A-U

5'-.....AGGAGAUUAUACCAUG ACUAGCAAAGGGAGAA... -3'

HL mit Stammlänge 7bp: -8,6 kcal/mol

C A  
A U  
U C  
U-A  
U-A  
C-G  
G-C  
U-A  
C-G  
A-U

5'-.....AGGAGAUUAUACCAUG ACUAGCAAAGGGAGAA... -3'

**Figur 2 (Fortsetzung)**

**HL mit Stammlänge 4bp:** -4,1 kcal/mol

U U  
A U  
A-U  
U-A  
C-G  
A-U

5'-.....AGGAGAUUAUACCAUG ACUAGCAAAGGAGAA... -3'

**HL mit Stammlänge 5bp:** -4,4 kcal/mol

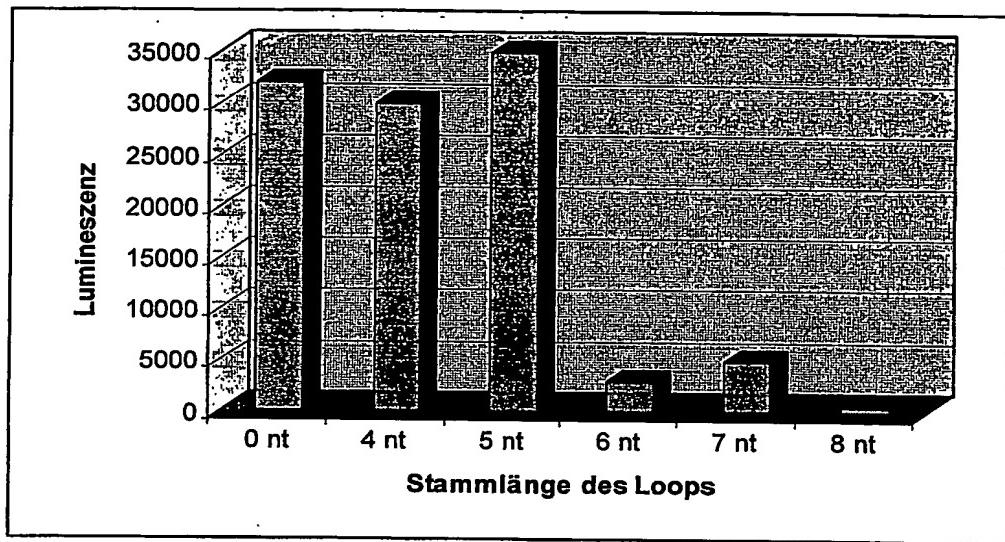
U A  
U U  
U-A  
G-C  
U-A  
C-G  
A-U

5'-.....AGGAGAUUAUACCAUG ACUAGCAAAGGAGAA... -3'

**HL mit Stammlänge 8bp:** -11,8 kcal/mol

U U  
G-C  
C-G  
A-U  
C-G  
G-C  
U-A  
C-G  
A-U

5'-.....AGGAGAUUAUACCAUG ACUAGCAAAGGAGAA... -3'

**Figur 3**

5/16

**Figur 4**

HL mit Stammlänge 8bp, 6 Basen nach Start-ATG:

G A  
U U  
G-C  
C-G  
A-U  
C-G  
G-C  
U-A  
C-G  
A-U

5'-.....AGGAGAUUAUACCAUGACUAGC AAAGGGAGAA... -3'

HL mit Stammlänge 8bp, 9 Basen nach Start-ATG:

G A  
U U  
G-C  
C-G  
A-U  
C-G  
G-C  
U-A  
C-G  
A-U

5'-.....AGGAGAUUAUACCAUGACUAGCAA GGAGAA... -3'

HL mit Stammlänge 8bp, 12 Basen nach Start - ATG:

G A  
U U  
G-C  
C-G  
A-U  
C-G  
G-C  
U-A  
C-G  
A-U

**Figur 4 (Fortsetzung)**

5'.....AGGAGAUAUACCAUGACUAGCAAAGGA    GAA... -3'

          G A  
          U    U  
          G-C  
          C-G  
          A-U  
          C-G  
          G-C  
          U-A  
          C-G  
          A-U

5'.....AGGAGAUAUACCAUGACUAGCAA A    GGAGAA... -3'

HL mit Stammlänge 8bp, 12 Basen nach Start - ATG:

          G A  
          U    U  
          G-C  
          C-G  
          A-U  
          C-G  
          G-C  
          U-A  
          C-G  
          A-U

5'.....AGGAGAUAUACCAUGACUAGCAAAGGA    GAA... -3'

**Figur 4 (Fortsetzung)**

- HL mit Stammlänge 8bp, 15 Basen nach Start-ATG:

G	A
U	U
G-C	
C-G	
A-U	
C-G	
G-C	
U-A	
C-G	
A-U	

5'-.....AGGAGAUAUACCAUGACUAGCAAAGGAGAA      GAA... -3'

- HL mit Stammlänge 8bp, 18 Basen nach Start-ATG:

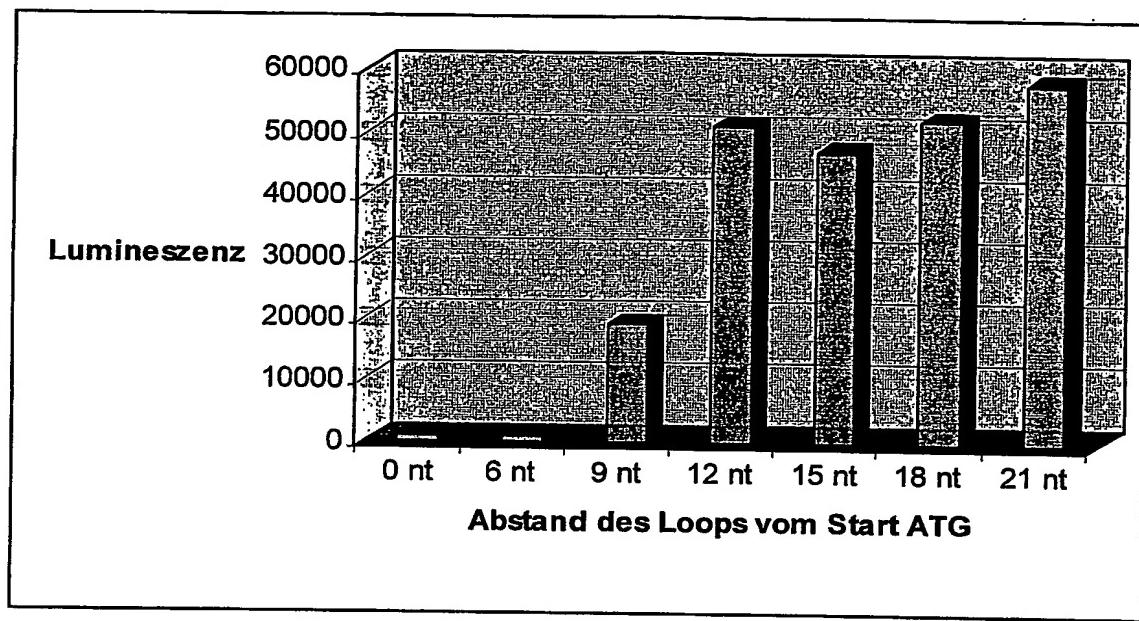
G	A
U	U
G-C	
C-G	
A-U	
C-G	
G-C	
U-A	
C-G	
A-U	

5'-.....AGGAGAUAUACCAUGACUAGCAAAGGAGAAGAA      CTT... -3'

- HL mit Stammlänge 8bp, 21 Basen nach Start-ATG:

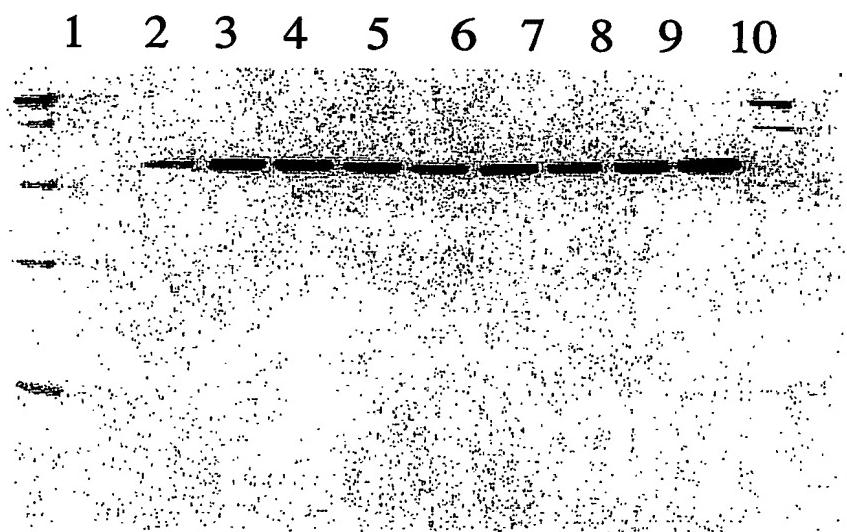
G	A
U	U
G-C	
C-G	
A-U	
C-G	
G-C	
U-A	
C-G	
A-U	

5'-.....AGGAGAUAUACCAUGACUAGCAAAGGAGAAGAACTT      TTC... -3'

**Figur 5**

**Figur 6**

**Spur 1-9= Mutante 9-1 Spur 1 entspricht der Wildtypsequenz**

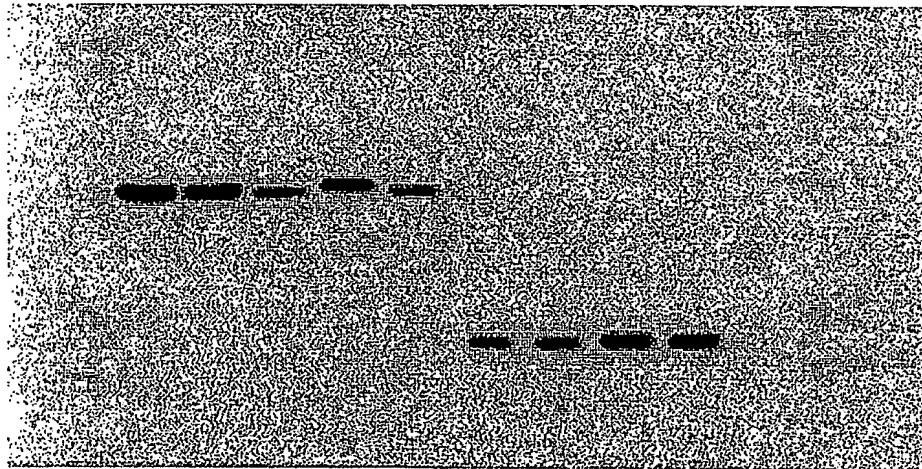


**Figur 7**

**GFP/1049**

**Spur 1,2= Mutante 1,2 GFP ohne Loop. Spur 3,4= Mutante 1,2 GFP mit Loop.**  
**Spur 5 entspricht der GFP Wildtypsequenz. Spur 6,7= Mutante 1,2 1049**  
**ohne Loop. Spur 8,9= Mutante 1,2 1049 mit Loop. Spur 10 entspricht der**  
**1049 Wildtypsequenz.**

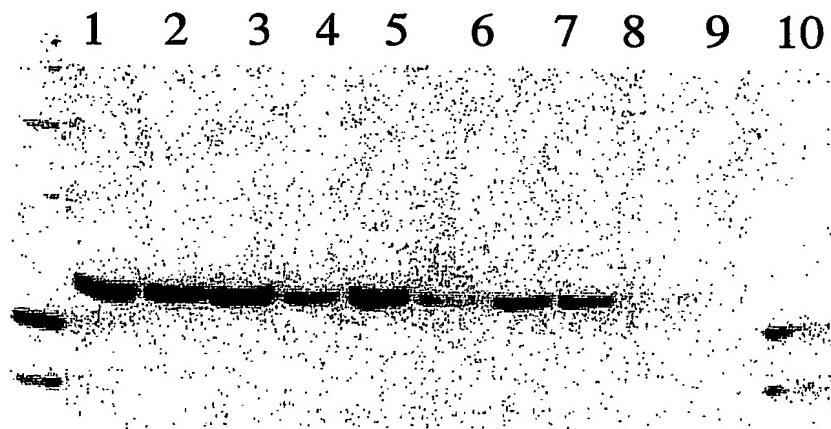
1 2 3 4 5 6 7 8 9 10



**Figur 8**

**Survivin**

**Spur 1-9= Mutante 1-9 Spur 10 entspricht der Wildtypsequenz**

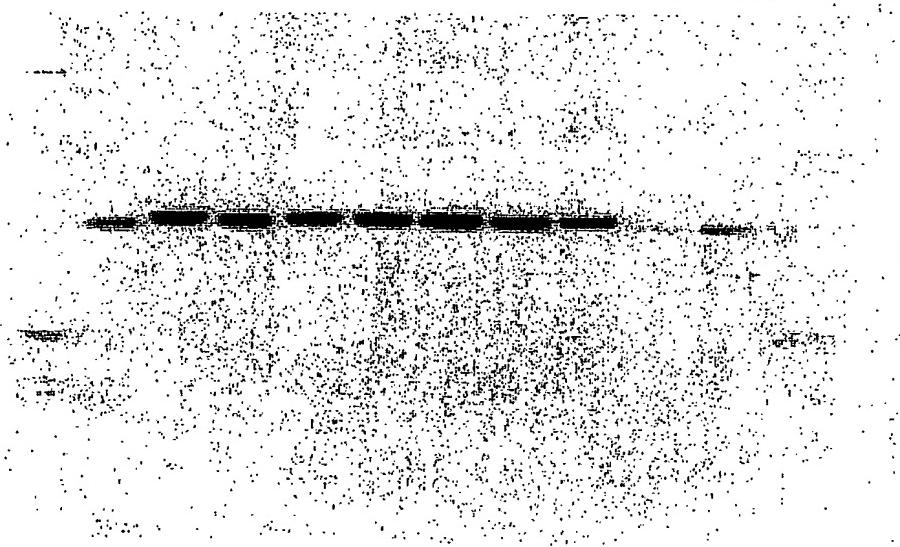


12/16

**Figur 9**

**GFP**  
**Spur 1-9 = Mutante 1-9 Spur 10 entspricht der Wildtypsequenz**

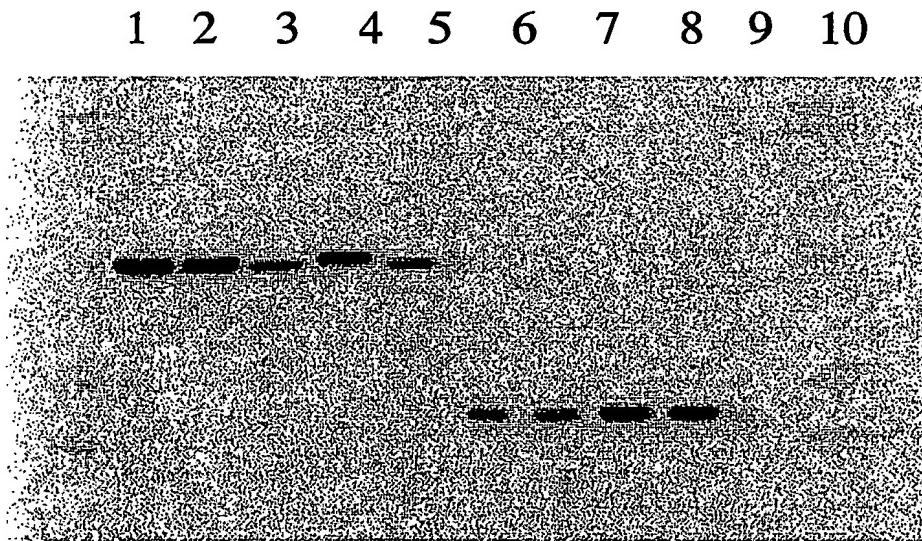
1 2 3 4 5 6 7 8 9 10



**Figur 10**

**GFP/1049**

**Spur 1,2= Mutante 1,2 GFP ohne Loop. Spur 3,4= Mutante 1,2 GFP mit Loop.  
Spur 5 entspricht der GFP Wildtypsequenz. Spur 6,7= Mutante 1,2 1049  
ohne Loop. Spur 8,9= Mutante 1,2 1049 mit Loop. Spur 10 entspricht der  
1049 Wildtypsequenz.**

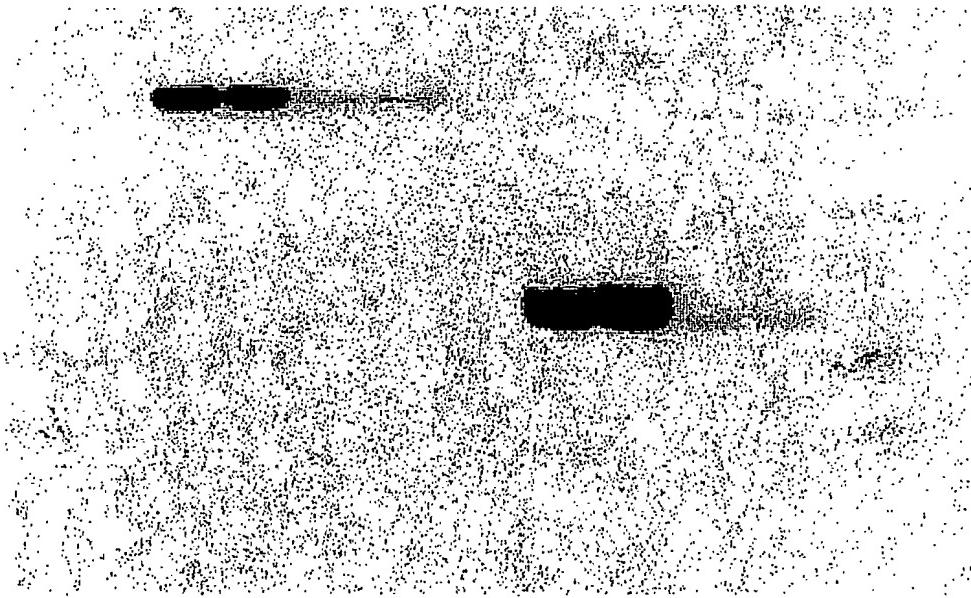


**Figur 11**

CIITA/Survivin

**Spur 1 entspricht der CIITA Wildtypsequenz. Spur 2,3= Mutante 1,2 CIITA mit Loop. Spur 4,5= Mutante 1,2 CIITA ohne Loop. Spur 6 entspricht der Survivin Wildtypsequenz. Spur 7,8= Mutante 1,2 Survivin mit Loop. Spur 9,10= Mutante 1,2 Survivin ohne Loop..**

1 2 3 4 5 6 7 8 9 10



**Figur 12**

Loop	Loop'
G A	G A
T T	A T
G-C	T-A
C-G	A-T
A-T	A-T
C-G	A-T
G-C	C-G
T-A	A-T
C-G	G-C
	A-T
	C-G-T-A

**Figur 13**

1049/Survivin/CIITA Loop und Loop'-Varianten

**Spur 1 1049 Mutante 1 Loop, Spur 2 Mutante 1 Loop', Spur 3 entspricht der 1049 Wildtypsequenz; Spur 4 Survivin Mutante 1 Loop, Spur 5 Mutante 1 Loop', Spur 6 entspricht der Survivin Wildtypsequenz; Spur 7 CIITA Mutante 1 Loop, Spur 8 Mutante 1 Loop', Spur 9 entspricht der CIITA Wildtypsequenz**

